University of Florida
Department of Economics
Fall 2021

Economic Analysis of Data
ECO4934

Instructor:	Ignacia Mercadal	**Time:**	M&W 9:35 – 11:30
Email:	imercadal@ufl.edu	**Classroom:**	TBD
Office Hours:	Thursdays 2:00–4:00	Sign up for an appointment here	
TA:	Ann Atwater	**TA email:**	annatwater@ufl.edu

Course Description and Objective

This is a course on economic data analysis using R. We will first learn to explore and visualize data. Then we will move to tools commonly used to find relationships between variables, where we will discuss the challenges arising when we try to establish causal relationships from observational data and how to address them. By the end of this class, you are expected to be able to explore and describe a data set using R, and think critically about when certain tools are able to separate correlation and causality in concrete examples. To put these skills into practice, students will work on a group research project in which they will pose a policy relevant question and work on answering using publicly available data.

Prerequisites

Principles of Microeconomics (ECO 2013), *Principles of Macroeconomics* (ECO 2023), and *Introduction to Statistics I* (STA 2023) or equivalent introductory statistics class. Students are expected to have basic knowledge of statistics and multivariate calculus. Please contact me if you are interested in the class and want to discuss whether you have the background to take this class. **Knowledge of R is not required**.

Class Format: HyFlex

This course is offered in the HyFlex format. You can choose to attend in person, or through the Zoom link provided here. Please note that this delivery modality could change as the public health situation evolves. The University, Department of Economics and I will continuously assess this situation. Regularly check your Canvas

inbox and UF email account for future information on class delivery. If you attend face-to-face, masks are expected in all economics classes. `https://coronavirus.ufl.edu/health-guidance/`.

Course Website

All information about the course will be posted on Canvas. Students are expected to check it regularly. All assignments must be submitted through Canvas.

Textbooks and Readings

The main textbooks for this class will be:

- Joshua D Angrist and Jörn-Steffen Pischke. *Mastering 'Metrics: The Path from Cause to Effect.* Princeton University Press, 2014

- Chester Ismay and Albert Young-Sun Kim. *Statistical inference via data science : a ModernDive into R and the Tidyverse.* 2020. Available here.

- Bruce Hansen's "Introduction to Econometrics", available here

The textbooks are mostly for reference since the class will be fast paced. I will point out to the specific sections as necessary, please send me an email if you want more guidance on how to use the books.

Group Project

You will work in groups of 2-3 students on a project. The goal will be to answer a policy relevant question using a publicly available dataset and analyzing it using R and Rstudio. I will guide you throughout the semester answering questions during office hours and giving you feedback on your intermediate deliveries. All submissions are expected to be done using RMarkdown and submitted to Canvas. Relevant dates:

- Identify question and data, timeline, and plan on team roles due on September 24
- Descriptive analysis and potential confounding factors, due on October 22
- Exploratory analysis, due on November 19
- In-class presentations on Dec 1, 6
- Final version due on December 10

Together with the final version, you will submit a 1 page summary of your contribution to the team, as well as a description of the contribution of the other members, and how well you worked as a team. Some good practices to work better as a group:

- Respond to emails within 24 hours.
- Arrive to meetings on time and prepared.
- Listen to others, and respond to comments with respect and honesty.
- Do your share.
- Do not underestimate the amount of time needed to complete each task.
- Any others on which you decide to agree.

Learning R

This is a hands-on class in which we will use R to analyze data throughout the semester. You will learn R primarily from practicing on your own, which is the only way to really learn. To help in this process, the final part of each class will be a lab in which you will work individually but will be able to ask questions to the professor. For this, you should bring your laptop every class.

Here are some resources that you might find useful. I encourage you to explore them independently:

- Google: "do X in R". Try variations of X until you find an answer. You will find many answers on https://stackoverflow.com/.
- Hadley Wickham and Garret Grolemund, *R for Data Science*. Available here
- Datacamp: An online resource to learn R and data science in general. You will have unlimited access during this semester. For this, sign up using this link.
- R Markdown resources: https://rmarkdown.rstudio.com

Problem Sets

There will be several individual problem sets during the semester. You can work in groups, but each student will submit an individual solution. Problem sets will be submitted as an RMarkdown knitted into pdf or html (you will learn how to do this in this class). The code should run without issues and generate the pdf or html output. The grader will not fix code bugs.

Late assignments will be discounted at 1 point per minute, which means that after 100 minutes you will not get any points. There will be no makeup assignments, but you will have some flexibility because only the best 5 out of 6 problem sets will count for your grade. Homework solutions will be posted on Canvas after they are due.

Exams

There will be two exams worth 20% each. If for **any reason** you cannot attend the first exam, the weight will be moved to the second exam. If you take both exams

and do better in the second one, only the second exam will count for your grade. I.e. the exam component of your grade is $\max\{0.2E1 + 0.2E2, 0.4E2\}$. If you cannot attend the second exam, please let me know as early as possible.

Cheating and Plagiarism:

Please review the Student Honor Code and Student Conduct Code. Violations will be subject to the University of Florida's Disciplinary Procedures.

Grades

- 10% of your grade will be based on participation, completion of DataCamp assignments, and performance in in-class labs.
- 25% of your grade will be performance on the best 5 out of 6 problem sets.
- 40% of your grade will be performance on the two exams (20% each).
- 25% of your grade will be a group project.

Regrade Requests Regrade requests must be typed and submitted in writing at the Professor's office within a week of receiving your graded problem set or exam back. You must state the exact reason for a regrade, either due to a miscalculation adding the points or an alternative but valid solution. "I think I deserve more points for this answer" is not an acceptable request and will not be considered. The whole assignment will be regraded, which means that *your grade can increase or decrease.*

Grading will be according to the following scheme:

Score	Grade
94 - 100	A
90 - 93	A-
87 - 89	B+
83 - 86	B
80 - 82	B-
77 - 79	C+
72 - 76	C
69 - 71	C-
66 - 68	D +
63 - 65	D
60 - 62	D-
0 - 59	E

COVID-19 Precautions

In response to COVID-19, the following practices are in place to maintain your learning environment, to enhance the safety of our in-classroom interactions, and to further the health and safety of ourselves, our neighbors, and our loved ones.

- If you are not vaccinated, get vaccinated. Vaccines are readily available at no cost and have been demonstrated to be safe and effective against the COVID-19 virus. Visit this link for details on where to get your shot, including options that do not require an appointment. Students who receive the first dose of the vaccine somewhere off-campus and/or outside of Gainesville can still receive their second dose on campus.

- You are expected to wear approved face coverings at all times during class and within buildings even if you are vaccinated. Please continue to follow healthy habits, including best practices like frequent hand washing. Following these practices is our responsibility as Gators.

- Sanitizing supplies are available in the classroom if you wish to wipe down your desks prior to sitting down and at the end of the class.

- Hand sanitizing stations will be located in every classroom.

- If you sick, stay home and self-quarantine. Please visit the UF Health Screen, Test & Protect website about next steps, retake the questionnaire and schedule your test for no sooner than 24 hours after your symptoms began. Please call your primary care provider if you are ill and need immediate care or the UF Student Health Care Center at 352-392-1161 (or email covid@shcc.ufl.edu) to be evaluated for testing and to receive further instructions about returning to campus. UF Health Screen, Test & Protect offers guidance when you are sick, have been exposed to someone who has tested positive or have tested positive yourself. Visit the UF Health Screen, Test & Protect website for more information.

- Course materials will be provided to you with an excused absence, and you will be given a reasonable amount of time to make up work.

- If you are withheld from campus by the Department of Health through Screen, Test & Protect you are not permitted to use any on campus facilities. Students attempting to attend campus activities when withheld from campus will be referred to the Dean of Students Office.

- Continue to regularly visit coronavirus.UFHealth.org and coronavirus.ufl.edu for up-to-date information about COVID-19 and vaccination.

Class Recordings

Please be aware of UF's policy on class recordings:

"Students are allowed to record video or audio of class lectures. However, the purposes for which these recordings may be used are strictly controlled. The only allowable purposes are (1) for personal educational use, (2) in connection with a complaint to the university, or (3) as evidence in, or in preparation for, a criminal or civil proceeding. All other purposes are prohibited. Specifically, students may not publish recorded lectures without the written consent of the instructor.

A "class lecture" is an educational presentation intended to inform or teach enrolled students about a particular subject, including any instructor-led discussions that form part of the presentation, and delivered by any instructor hired or appointed by the University, or by a guest instructor, as part of a University of Florida course. A class lecture does not include lab sessions, student presentations, clinical presentations such as patient history, academic exercises involving solely student participation, assessments (quizzes, tests, exams), field trips, private conversations between students in the class or between a student and the faculty or lecturer during a class session.

Publication without permission of the instructor is prohibited. To "publish" means to share, transmit, circulate, distribute, or provide access to a recording, regardless of format or medium, to another person (or persons), including but not limited to another student within the same class section. Additionally, a recording, or transcript of a recording, is considered published if it is posted on or uploaded to, in whole or in part, any media platform, including but not limited to social media, book, magazine, newspaper, leaflet, or third party note/tutoring services. A student who publishes a recording without written consent may be subject to a civil cause of action instituted by a person injured by the publication and/or discipline under UF Regulation 4.040 Student Honor Code and Student Conduct Code."

Class Content

1. Introduction to Economic Analysis of Data and to R

   - *Mastering Metrics*, Introduction
   - R and Rstudio installation and basics

2. Data Description and Visualization

   - *Modern Dive*, Chapters 2, 3, 4

3. Simple Linear Regression

   - *Modern Dive* Ch. 5

4. Introduction to Causality

   - *Mastering Metrics* Ch.1

5. Multiple Linear Regression

   - Mastering Metrics Ch. 2
   - *Modern Dive*, Ch.6

6. Sampling

   - *Modern Dive*, Ch.7

7. Confidence Interval and Hypothesis Testing

   - *Modern Dive*, Ch.8 and 9

8. Regression Inference

   - *Modern Dive*, Ch.10
   - SPo Ch. 6

9. Panel Data and Differences-in–Differences

   - *Mastering Metrics*, Ch. 5

10. Topics: IV or RDD (time dependent)

    - *Mastering Metrics*, Ch. 3,4

Tentative Schedule:

MONDAY	WEDNESDAY	FRIDAY
8/ 23 Syllabus and R	8/25 Bring laptop. Statistics Review. Intro to R and Rstudio.	8/27
8/30	9/ 1	9/3 **11am** Group members due Problem Set 1
9/6 **Labor Day**	9/8	9/10
9/13	9/15	9/17 Problem Set 2
9/20	9/22	9/24 Project Part 1
9/27	9/29	10/ 1 Problem Set 3
10/4 Review Session	10/6 **First Exam**	10/8 **Homecoming**
10/11	10/13	10/15
10/18	10/20	10/22 Project Part 2
10/25	10/27 Problem Set 4	10/29
11/ 1	11/3	11/5

MONDAY	WEDNESDAY	FRIDAY
11/8	11/10 Problem Set 5	11/12
11/15	11/17	11/19 Project Part 3
11/22 Problem Set 6	11/24 **Thanksgiving Break**	11/26 **Thanksgiving Break**
11/29 Review Session	12/ 1 Student Presentations	12/3
12/6 Student Presentations	12/8 **Second Exam**	12/10
12/13 Final version of the project	12/15	12/17